An Introduction to Neuro-AI Computing

Fukumi KOZATO

This paper introduces symbolic AI and connectionism in general, and clarifies their pros and cons. Based on those, the importance of integrating the two paradigms to model sophisticated cognitive systems is discussed.

1. Connectionism and Human Cognition

Connectionism is a relatively new approach to investigating the human cognitive mechanism. It focuses on computational architecture and dynamics to study the massive parallelism that the human brain possesses, and is motivated by the following idea [1]:

Any cognitive activity of human beings is ultimately supported by the massively parallel operation of neurons, even if it appears as a serial process. Therefore, investigating massive parallelism may reveal the mechanism of the human cognition.

Cerebrophysiological research has made rapid progress in recent years and neurodynamics is being elucidated [2] [3]. Nowadays, we know the structural characteristics of the brain and the dynamics of individual neurons quite well. For instance, we know how a neuron looks, how it connects to others, how it exchanges signals with others, and so forth. However, we have very little knowledge of the dynamic features of the brain as a whole. We do not know how information is kept in the brain or processed at all. This is especially the case when we consider high level human cognitive activities. By any means, the human brain cannot be examined thoroughly as long as it is alive and working.

Connectionism is said to be "neurally inspired modelling" [4], but the modelling is not an exact copy of the biological brain system. Although its architecture is often disigned on the basis of the brain structure to gain the massive parallelism, the method of information representation and processing may be developed just for a particular architecture to implement a particular human cognitive activity in it. Consequently, most connectionist models seem to have very little similarity to actual human brain dynamics. After all, we only know the hardware but not the software.

電子工学科
1993年9月3日受理

1. 1. Definition

Connectionism comes in many different forms, and the definition is wide open to debate. As the research field has grown, all sorts of models have been developed. Among the various definitions, however, the most traditional one may be expressed as follows [4], [5], and [6] :

(1) A connectionist architecture is a network that consists of a great number of processing units and connection links between them.
(2) Each unit has only very simple structure and function, and communicates with others by exchanging simple messages.
(3) Information is stored by the pattern of interconnections among units.
(4) Units work massively in parallel with no centralised controller.

1. 2. Structure

The structure of connectionist models in basically divided into two groups. One is *feedforward* type networks (see Figure 1) in which units are arranged in layers and no unit connects to the others within a layer. The flow of activation signals is only in one direction so that activation signals provided at the input layer are sent over to the next layer towards the output layer. NET talk [7], which has been taught to pronounce text is a well known application using a feedforward network, and RAAM networks [8] which is a recursive auto-associative memory for encoding and decoding syntactic trees is a recent example of a feedforward network. Some of the networks in this category may have feedback connections between particular layers, e.g., recurrent networks [9] and [10]. The other may be called *feed-any* type networks (see Figure 2) in which units may connect to any other unit of the network possibly including itself. The most popular example seems to be the Hopfield network [11].

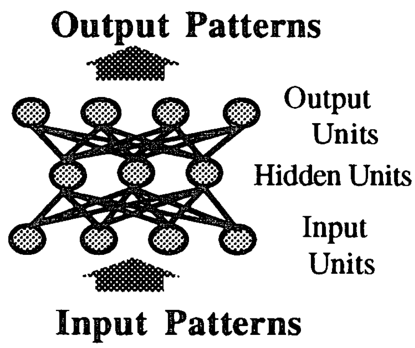

Figure 1. A Feedforward Network

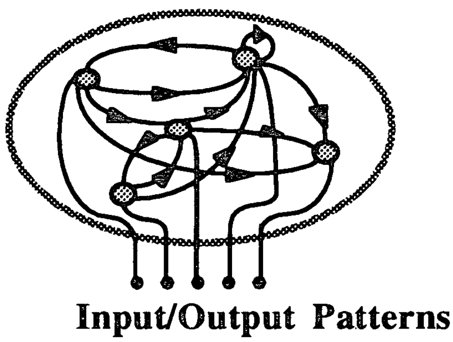

Figure 2. A Feed-and Network

1.3. Dynamics

The dynamics of connectionist networks do not stand alone but depend on the network structure. There are a number of variations from well-known to very specific ones. However, they can be roughly categorised into two groups.

We call the first group *cost minimisation*. A typical example is that of the Hopfield network. The principle of cost minimisation is to set up a cost function that represents the state of a connectionist network and minimise the cost. Low cost indicates a highly stable state, whereas high cost indicates low stability. Cost minimisation proceeds basically as follows:

Step 1: A network is given an initial state (input information) as a unit activation pattern p. The initial cost C of the network state $S_{net}(p)$ is given by $C=f(S_{net}(p))$ where f is the cost function of the network.

Step 2: The network moves to a particular stable state $S_{net}(p')$ by changing the initial activation pattern p to p', where $C'=f(S_{net}(p'))<C$.

Step 3: A certain network state $S_{net}(p'')$ given by a unit activation pattern p'' expresses the final stable state (output information), only if $C''=f(S_{net}(p''))$ indicates the minimum cost.

We can call the second group *spreading activation*. For cost minimisation, the state of a whole network is taken into account, whereas for spreading activation the state of individual units is looked at. This method may be applied to a quite wide range of connectionist networks. Both feedforward and feed-any type networks can employ this dynamics. Spreading activation proceeds as follows:

Step 1: The input units of a network are given an initial unit activation pattern p (input information) that consists of a set of activation signals a1, a2.., a_n for each input unit.

Step 2: The activation signals are propagated from the input units to the adjacent ones through connections, one after another. The state of a particular unit $S_j(a_j)$ is given by

$$S_j(a_j)=h\left(\sum_{k=0}^{l} g(S_k(a_k))\right)$$

where $S_k(a_k)$ is the state of unit k that sends out an activation signal to the unit j, l is the number of units that send out an activation signal to the unit j, and g and h are propagation functions.

Step 3: The propagation process stops at particular units when a certain terminal condition has been satisfied, e.g., running out of the activation signals by a decay function or reaching the output units. The output information can be expressed by the unit activation pattern at the output site or the level of activation where the propagation has stopped.

2. Symbolic Artificial Intelligence and Human Cognition

Sypbolic arificial Intelligence (AI) is the study how to make conventional computers do things at which people are better [12], and provides useful techniques for modelling human cognitive activities or for solving complex problems. Symbolic AI attaches importance especially to the investigation of the serial flow of human cognition and aims to simulate it by manipulationg symbols on conventional computers. This contrasts with connectionism that is based on the study of the massive parallelism that the human brain possesses and aims to develop new architecture.

It is generally recognised that the first formal proposal of symbolic AI as we know it today was at

Dartmouth Conference in 1956 [13]. One of the conference organisers John McCarthey addressed that

> *The study (of AI) is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.*

In those days, there were mainly two approaches for modelling human intelligence. One was to simulate the massively parallel information processing mechanism of the human brain itself, and the other was to model a serial flow of human conscious cognition by symbols and manipulate it in a conventional computer. The Former idea was succeeded by connectionism, whereas the latter became the base of symbolic AI.

One of the most flourishing and practical applications of symbolic AI is *expert systems* that handle human intelligent activities in a specific field. They perform rule-based problem solving, which is one of the core techniques of symbolic AI, so as to require knowledge in the form of rules and axioms for a particular domain within which certain problems are solved. For expert systems, the knowledge domain is restricted to a certain expert field for which know-how to solve problems has been concretely established. The medical diagnosis system MYCIN [14] and the mineral exploration advisor PROSPECTOR [15] are both well-known expert systems. An expert system consists of three components, i.e., a knowledge base, an inference engine, and a user interface, and performs inferences as serial deduction processes. A knowledge base contains the domain knowledge consisting of production rules (*IF-THEN* rules) and facts (axioms). An inference engine operates a knowledge base to manipulate the knowledge to draw inferences, and a user interface interacts with the user to ask and answer queries.

The proposal of expert systems has been generally accepted in a favourable mood as a practically useful application of symbolic AI technology, even though it focuses on a very limited view of human cognition. However, some serious problems of symbolic AI have also been revealed. One of them is that vague knowledge cannot be easily represented in symbolic form, and another is that serial inference may result in a combinatorial explosion [1].

3. Combining Connectionism and Symbolic AI

There are a variety of arguments on connectionist and symbolic AI approaches for modelling human coginition. In this section, some of the significant arguments are introduced to clarify the weakness of both appoaches. Although it is obvious from the arguments that these two paradigms are conflicting with each other, they can be usefully combined by utilising the features of each technique complementarily.

3.1. Connectionism vs. Symbolic Artificial Intelligence

Although connectionism may be quite successful for modelling low level perceptual functions, e.g., visual or speech pattern recognition, it may not be so successful for the modelling of human conscious or higher cognitive activities, e.g., reasoning or planning. This is because, in the parallel architecture of connectionist models, the former can be much more simply achieved than the latter. Besides that, some connectionists may still believe that the key to understanding the principal mechanisms of human cognition can be found only in the extension of the study of human unconscious cognition.

Until recently, the modelling of human cognition has been attempted mainly using methods based on symbolic AI, or in other words the serial symbolic information processing paradigm. However, connectionists argue against such models. They point out important missing factors that the human brain always possesses, and claim that their absence makes symbolic AI look very unrealistic. For example, [5] summarises those missing factors as (1) efficient search techniques, (2) pattern recognition ability, and (3) representation of non-symbolic information, and claim that

> *The traditional AI approach so knowledge and recoginition problems is to use ever more complex and clever strategies to reduce the need for excessive search and computation.*

[16] doubts the ability of symbolic AI techniques for modelling human cognition from a different point of view. His question is based on what he calls the *symbol grounding problem.* He argues that

> *the (semantic) interpretation (of symbols and symbol manipulation) will not be intrinsic to the symbol system itself : it will be parasitic on the fact that they have meaning for us. Hence, if the meanings of symbols in a symbol system are extrinsic rather than intrinsic like the meanings in our heads, then they are not a viable model for the meanings in our heads : Cognition cannot be just symbol maniqulation.*

[17] also discusses a similar point as follows:

> *The purely symbolic paradigm of much of AI suffers from not being grounded in perception. ....Higher level cognitive activities, such as understanding speech or recognising images, which do not attempt to use low-level models may be doomed to failure.*

On the other hand, connectionism is still controversial in the symbolic AI society, and there is obvious criticism to connectionism. That is to say, connectionism cannot handle human conscious cognition properly, because it must be modelled on the manipulation of symbolic information and connectionism may not be capable of handling symbol systems. From this point of view. [18] pursues a strong argument against connectionism, i.e.,

> *Connectionism can not properly model human conscious cognition because it does not provide semantic compositionality nor syntactic systematicity to the model.*

Such an argument is mainly raised by the cognitive scientists who enthusiastically favour the symbolic AI approach. However, [19] comments on this claim as follows:

The paper ([18] written by Fodor and Pylyshyn) may be written too early concerning yet an incomplete and unexplored theory.

Connectionist cognitive scientists complain that [18] has set a very strict definition of connectionist models, and asserted the following conditions in connectionism:

(1) All connectionist representations are atomic.
(2) Connectionist processing is *association* which is sensitive only to statistics, not structure.

On this point, there are a number of counterarguments from connectionist cognitive scientists. For example, [20] attempts to refute [18] as follows:

Fodor and Pylyshyn are simply mistaken in their claim that connectionist mental states lack the necessary consitituent structure, and that the basis of this mistake is a failure to appreciate the significance of distributed representations in connectionist models.

[21] and [22] also support this argument showing that connectionist models can employ compositionally structured representations in a manner quite different from the classical symbol processing approach. [21] points out that

Fodor and Pylyshyn are implicitly discussing only one type of compositionality, i.e., concatenative composition,

and suggests another compositionality that is available to connectionist modelling. He says

we (connectionists) have functional compositionality when there are general, effective and reliable processes for producing an expression given its constituents, and decomposing the expression back into these constituents.

3. 2. Hybrid Modelling

Because of the disagreement of two parabigms described in **3.1**, connectionism and symbolic AI are often viewed as rivals. However, that idea may not be very productive, because contrary technologies can often work as useful complements to each other. [17] explains this based on the following statement and claims the necessity of developing hybrid connectionist/symbolic systems.

....there is a clear and present need for developing systems which can perform both 'perceptual' and 'cognitive' tasks... To provide 'low-level' perceptual functionality as well as demonstrating high-level cognitive abilities we need to capture the best features of

current connectionist and symbolic techniques.

In order to realise this, there should be four ways [17].

(1) to figure out a methodology for getting traditional AI systems to handle low-level functions,
(2) to figure out a methodology for getting connectionist systems to handle high-level symbol processing tasks,
(3) to work out a new paradigm for delivering intelligent behaviour, and
(4) to take the current connectionist systems and the current generation of AI systems and produce hybrid systems exploiting the strengths of each.

He emphasises the importance of investigating the fourth method: he simply says

> *building hybrid models requires no major developments, but rather the linking of current technologies.....From a purely applied perspective, we see a fine reason to pursue the building of hydrid models.*

But this should not be the only reason why hybrid models need to be investigated.

[23] provides a strong argument why hybrid modelling of intelligence is necessary. They claim that a comprehensive theory of intelligence will require both connectionist and symbolic approaches, because their relative strengths and weaknesses suggest a natural division of labour within a hybrid system. They point out strength and weakness of connectionist models as follows:

> *Non-symbolic connectionist systems (e.g., [24] has shown plausible connectionist modelling of the stages that children pass through in acquiring the ability to from past tense), provide a more satisfactory explanation of certain aspects of cognition than symbol systems do, whereas the lack of universality of non-symbolic connectionist systems (because of the bounded storage capacity) supports the claim that some symbolic processing is necessary for intelligence.*

[1] supports those arguments in a general statement:

> *All creatures where the information processing mechanism has been developed as neural network are surely the realisation of connectionism. Yet, it is wrong to ignore the other fact that the serial information processing to handle symbolic information was absolutely necessary for the development of our intelligence in the evolution from primitive creatures to the human beings.*

The above discussion implies at least the fact that the hybrid systems of symbolic AI and connectionist models is worth investigating. Connectionist models may be particularly superior in dealing with non-structural information expressed non-symbolically, whereas symbolic AI can be successful in representing and processing highly structured information.

Although definition of hybrid modelling is varied, the least necessary property of hybrid models should be mixed aspects of classical, symbolic processing and connectionist, sub-symbolic processing. [25] suggests two distinctive views of hybrid systems, namely the *strong* view and the *weak* view. From his discussion, a strong hybrid system must consist of two (or more) virtual machines, i.e., a classical symbol system and a connectionist system, and a weak hybrid system requires only one of those, i.e., mostly a connectionist system. The specific concerns of the researchers on strong hybrid systems are

(1) What role should each form of computation play in the hybrid systems?
(2) What should the interface be between the two forms of computation?

Whereas the primary concerns in a weak hybrid system are

(1) What properties of each form of computation are important?
(2) How can one form of computation be made to realise the important properties of the other form?

4. Conclusion

The researchers on the symbolic AI side claim that connectionism cannot handle human conscious cognition, because it must be modelled in the manipulation of symbolic information and connectionism may not be capable of handling symbol systems. On the other hand, those on the connectionism side argue that the key to understanding the principle of the mechanisms of human cognition can be found in the extension of the human unconscious cognition study for which symbolic AI is not adequate. However, such a critical attitude, excluding another technology for solving a similar problem, may not be very productive.

Symbolic AI is regarded as highly successful in representing and processing knowledge as long as information can be replaced by a formal language, but it is limited by the conventional formality of symbols and sequential information processing. Connectionist models are quite suitable for handling unstructured primitive data or non-symbolic information rather than high level information because the connectionist structure is basically a regular arrangement of simple processing units and the operation is organised as a collection of uncontrolled actions of autonomous units. Thus, adopting the useful part of the other technology can be beneficial for constructing much more sophisticated cognitive systems, because both technologies can be considered as complementary to each other, and thus investigating hybrid models should be a very positive idea.

References

[1] Amari, S., "Brain and Computer" in *Neural Networks and Connectionism* written in Japanese, chap. 1,

University Press, Tokyo, 1989.

[2] Ito, M., *Design of the Brain* written in Japanese (Japanese tjtle: Nou-no Sekkeizu), Chuokouronsha (pub.), Tokyo, 1980.

[3] Sakata, H., Anzai, U. and Amari, S., *Neuro-science at Present* written in Japanese, Chuko-Shuppan, Tokyo, 1987.

[4] Arbib, M. A., "Learning Networks" in *Brains, machines, and mathematics : 2nd edition,* chap. 5, Springer-Verlag, New York, 1987.

[5] Fahlman, S. E. and Hinton, G. E., Connectionist Architecture for Artificial Intelligence, IEEE: Computer, pp. 100-109, January 1987.

[6] Addanki, S., "Connectionism" in *Encyclopedia of Artificial Intelligence,* vol. I, pp.268-274, Shapiro, S. C. (ed.), Wiley-Interscience, New York, 1992.

[7] Sejnowski, T. J. and Rosenberg, C. R., Parallel Networks that Learn to Pronounce English Text. Complex Systems, vol. 1, no. 1, pp.145-168, 1987.

[8] Pollack, J., Recursive Auto-Associative Memory: Devising Compositional Distributed Repressentations. Proc. of the Twelfth Annual Conference of Cognitive Science Society, pp. 33-39, 1990.

[9] Jordan, M. I., Serial Order: A Parallel Distributed Processing Approach. Technical Report no. 8604, San Diego: University of California, Institute for Cognitive Science, 1986.

[10] Elman, J. L., Finding Structure in Time. Cognitive Science, vol. 14, pp.179-211, 1990.

[11] Hopfield, J. J., Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. of National Academy of Science of USA, vol. 79, pp.2554-2558, 1982.

[12] Rich, E., *Artificial Intelligence,* McGraw-Hill, New York, 1983.

[13] Charniak, E. and McDermott, D., "AI and Internal Representation" in *INTRODUCTION TO Artificial Intelligence,* chap. 1, Addison-Wesley Publishing Company, Massachusetts, 1985.

[14] Shortliffe, E. H., *Computer-based Medical Consultations : MYCIN,* Elsevier, New York, 1976.

[15] Duda, R. O., Gaschnig, J. G, and Hart, P. E., "Model Design in the Prospector Consultant System for Mineral Exploration" in *Expert Systems in the Microelectronic Age,* pp.153-167, Edinburgh University Press, Michie, D. (ed), Edinburgh, 1980.

[16] Harnard, S., The Symbol Grounding Problem. Presented at CNLS Conference on Emergent Computation, Session on "Emergencce of Symbols", Submitted to Physica, Los Alamos, USA, 1989.

[17] Hendler, J. A., "Developing Hybrid Symbolic/Connectionist Models" in *High Level Connectionist Models,* Barnden, J. A. and Pollack, J. B. (eds), vol. 1, chap. 7, Ablex Publishing Corporation, New Jersey, 1991.

[18] Fodor, J. A. and Pylyshyn, Z. W., Connectionism and cognitive architecture: a critical analysis. Cognition, vol.28, pp.7-71, 1988.

[19] Aleksander, I. and Morton, H., COGNITIVE AUTOMATA: A Refutation of the Neuron/Symbol Divide (provisional title) in *Neurons and Symbols,* chap. 7, Chapman & Hall, London, to be published in 1993.

[20] Smolensky, The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn. The Southern Journal of Philosophy, vol. XXVI, Supplement, 1987.

[21] Gelder, T. V., "Compositionality: A Connectionist Variation on a Classical Theme", Cognitive Science, vol. 14, 66.355-384, 1990.

[22] Chalmers, D. J., Why Fodor and Pylyshyn Were Wrong: The Simplest Refutation. Proc. of the Eleventh Annual Conference of the Cognitive Science Society, pp.340-347, 1989.

[23] Derthick, M. and Plaut, D. C., Is Distributed Connectionism Compatible with the Physical Symbol System Hypothesis?, Proc. of the 8th Annual Conference of the Cognitive Science Society, pp.639-644, 1986.

[24] Rumelhart, D. E. and McClelland, J. L., "On learning the Past Tenses of English Verbs" in *Parallel Distributed Processing : Explorations in the Microstructure of Cognition,* Rumelhart, D. E. and McClelland, J. L. (eds), vol.2, chap. 18, MIT press, 1986.

[25] Stark, R. J., Does Hybrid Mean More Than One? Report on the JCI Workshop on Hybrid Models. AISB

Quarterly Newsletter of the Society for the Study of Artificial Intelligence & Simulation of Behaviour, no. 78, pp.8-9, 1991.